

IN THE CLAIMS:

The text of all pending claims, (including withdrawn claims) is set forth below. Cancelled and not entered claims are indicated with claim number and status only. The claims as listed below show added text with underlining and deleted text with ~~strikethrough~~. The status of each claim is indicated with one of (original), (currently amended), (cancelled), (withdrawn), (new), (previously presented), or (not entered).

Please AMEND claims 1, 13, 15, 17, 18, 20, 29, 31 and 32 and ADD new claims 33-35 in accordance with the following:

B1 1. (currently amended) A document searching apparatus for searching a document group having a link relation for a document, comprising:

a link importance assigning unit weighting the link relation and assigning link importance which indicates importance of the document based on the weighted link relation to the document; and

an accessing unit accessing the document based on the link importance, and

wherein said link importance assigning unit includes:

a URL similarity calculating unit calculating a URL similarity that is a text similarity of character strings of URLs (Uniform Resource Locators) that represent the location of the documents and that is an appearance of written characters of URLs,

wherein said link importance assigning unit calculates the link importance based on ~~the an inverse~~ URL similarity and the link relation of the document, so that a link weight increases as URL similarity decreases.

2. (cancelled)

3. (original) The document searching apparatus as set forth in claim 1, further comprising:

a keyword extracting unit extracting text parts from the documents and extracting a keyword from the document contents.

4. (original) The document searching apparatus as set forth in claim 3, wherein said keyword extracting unit calculates an occurrence frequency of the keyword in the document, and

wherein said keyword extracting unit further comprises:

B1
a keyword - document correlation calculating unit calculating the correlation of the keyword and the document based on the link importance and the occurrence frequency of the keyword.

5. (original) The document searching apparatus as set forth in claim 4, further comprising:

a monitoring unit monitoring accesses from a user and generating an access log, and wherein said keyword - document correlation calculating unit calculates the correlation based on the keyword occurrence frequency, the link importance, and the access log.

6. (original) The document searching apparatus as set forth in claim 4, further comprising:

a document type determining unit determining a document type of the document based on the URL similarity, the number of links from the document, and the number of links to the document,

wherein said keyword - document correlation calculating unit selects the document based on the document type and calculates the correlation for the selected document.

7. (previously amended) The document searching apparatus as set forth in claim 4, further comprising:

an index creating unit creating an index for accessing the document corresponding to pronunciation characters or spelling of the extracted keyword.

8. (original) The document searching apparatus as set forth in claim 7, further comprising:

a selecting unit allowing the user to select a portion of the pronunciation characters or spelling of the keyword,

wherein said index creating unit places less than a predetermined number of documents highly correlated with the keyword in the index based on the correlation calculated by said keyword - document correlation calculating unit, and

wherein said accessing unit accesses the document based on the selected keyword.

9. (previously amended) The document searching apparatus as set forth in claim 1, further comprising:

a collecting unit collecting the document from a network.

10. (original) The document searching apparatus as set forth in claim 1, wherein said link importance assigning unit causes the weight of the link relation between the documents with a high URL similarity to be decreased.

11. (previously amended) The document searching apparatus as set forth in claim 1, wherein said link importance assigning unit causes the document that is linked from an important document and whose URL similarity is low to be important.

12. (previously amended) The document searching apparatus as set forth in claim 1, wherein said link importance assigning unit causes the importance of a document linked from many documents whose URL similarity are high to be decreased.

13. (currently amended) A document searching apparatus for searching a document group having a link relation for a document, comprising:~~The document searching apparatus as set forth in claim 1,~~

a link importance assigning unit weighting the link relation and assigning link importance which indicates importance of the document based on the weighted link relation to the document; and

an accessing unit accessing the document based on the link importance, and wherein said link importance assigning unit includes:

a URL similarity calculating unit calculating a URL similarity that is a similarity of character strings of URLs (Uniform Resource Locators) that represent the location of the documents and that is an appearance of written characters of URLs,

wherein said link importance assigning unit calculates the link importance based on the URL similarity and the link relation of the document, and

wherein the link importance of each document is defined as a solution of the following simultaneous linear equation (1), assuming that C_q is constant (the lower limit of the importance that depends on each page) for each $p \in DOC$ and that when a page p is linked to a page q , the link weight $lw(p, q)$ is defined by the formula (2):

$$Wq = Cq + \sum_{p \in \text{Refed}(q)} Wp * lw(p, q) \quad \dots (1)$$

$$lw(p, q) = diff(p, q) / \sum_{i \in Ref(p)} diff(p, i) = \frac{1}{sim(p, q) \sum_{i \in Ref(p)} \frac{1}{sim(p, i)}}$$

... (2)

where DOC = {p1, p2, ..., pN} is a set of documents calculated for the link importance; Wp is the link importance of the page p; Ref(p) is a set of pages linked from the page p; Refed(p) is a set of pages linking to the page p; sim(p, q) is the URL similarity of the pages p and q; diff(p, q) = 1/sim(p, q) is the difference.

14. (previously amended) The document searching apparatus as set forth in claim 1,

wherein the URL similarity is determined based on characters of a URL containing a server address.

15. (currently amended) A document index creating apparatus for creating an index of a document group having a link relation, comprising:

a link importance assigning unit assigning a link importance to the document based on the link relation;

a keyword extracting unit extracting a keyword from the document;

an index creating unit creating an index for accessing the keyword based on pronunciation characters or spelling of the extracted keyword; and

an accessing unit accessing document assigned the link importance corresponding to the keyword when the pronunciation characters or spelling of the keyword are selected from the index, and

wherein said link importance assigning unit includes:

a URL similarity calculating unit calculating a URL similarity that is a text similarity of character strings of URLs (Uniform Resource Locators) that represent the location of the documents and that is an appearance of written characters of URLs,

wherein said link importance assigning unit calculates the link importance based on the an inverse URL similarity and the link relation of the document, so that a link weight increases as URL similarity decreases.

16. (cancelled)

B1 17. (currently amended) A document index creating apparatus for creating an index of a document group having a link relation, comprising:

a link importance assigning unit assigning a link importance to the document depending on whether or not URLs of the documents are similar;

a keyword extracting unit extracting a keyword from the document; and

an index creating unit creating an index for accessing the document corresponding to pronunciation characters or spelling of the extracted keyword based on the link importance, and

wherein said link importance assigning unit includes:

a URL similarity calculating unit calculating a URL similarity that is a text similarity of character strings of URLs (Uniform Resource Locators) that represent the location of the documents and that is an appearance of written characters of URLs,

wherein said link importance assigning unit calculates the link importance based on the an inverse URL similarity and the link relation of the document, so that a link weight increases as URL similarity decreases.

18. (currently amended) A link list creating system for creating a link list for a document group having a link relation, comprising:

a collecting unit collecting the documents from a network;

a link importance assigning unit assigning a link importance of the document as an importance calculated based on the link relation to the document;

a URL character string determining unit determining a URL having a particular characteristic of a character string from the documents;

an index creating unit creating a link list for listing less than a predetermined number of links to the documents based on the link importance and the particular characteristic of the character string of the URL; and

a document type determining unit determining a document type based on a URL similarity representing a text similarity between character strings of URLs (Uniform Resource Locators) of the documents and being an appearance of written characters of URLs, the number of links to the document, and the number of links from the document, and

wherein said index creating unit selects the document based on the document type and creates a link list of the selected document, and

wherein said link importance assigning unit calculates the link importance based on an inverse URL similarity and the link relation of the document, so that a link weight increases as URL similarity decreases.

B1
19. (cancelled)

20. (currently amended) A document searching method for searching a document group having a link relation for a document, comprising:

assigning a link importance as an importance of the document calculated with weighting the link relation to the document, comprising:

calculating a URL similarity that is a text similarity of character strings of URLs (Uniform Resource Locators) that represent the location of the documents and that is an appearance of written characters of URLs; and

calculating the link importance based on the URL similarity and the link relation of the document with said link importance being based on an inverse URL similarity and the link relation of the document, so that a link weight increases as URL similarity decreases;
and

accessing the document based on the link importance.

21. (cancelled)

22. (original) The document searching method as set forth in claim 20, further comprising:

extracting a keyword from the document.

23. (original) The document searching method as set forth in claim 20, further comprising:

calculating an occurrence frequency of the keyword in the document, and

calculating the correlation of the keyword and the document based on the link importance and the occurrence frequency of the keyword.

24. (previously amended) The document searching method as set forth in claim 23, further comprising:

monitoring accesses from a user and generating an access log; and

calculating the correlation based on the keyword occurrence frequency, the link importance, and the access log.

B1 25. (original) The document searching method as set forth in claim 23, further comprising:

determining a document type of the document based on the URL similarity, the number of links to the document, and the number of links from the document; and

selecting the document based on the document type and calculating the correlation of the selected document.

26. (previously amended) The document searching method as set forth in claim 22, further comprising:

creating an index for accessing the document corresponding to pronunciation characters or spelling of the extracted keyword.

27. (previously amended) The document searching method as set forth in claim 26, further comprising:

placing less than a predetermined number of documents which are correlated with the keyword in the index;

selecting a portion of the pronunciation characters or spelling of the keyword; and

accessing the document corresponding to the selected portion of the pronunciation characters or spelling of the selected keyword.

28. (previously amended) The document searching method as set forth in claim 20, further comprising:

collecting the document from a network.

29. (currently amended) A link list creating method for creating a link list for a document group having a link relation, comprising:

collecting the document from a network;

assigning a link importance which indicates importance of the document to the document based on the link relation;

determining a URL having a particular characteristic of a character string from the URLs of each document;

creating a link list for listing less than a predetermined number of links to the document based on the link importance and the particular characteristic of the character string of the URL

B1
determining a document type based on a URL similarity that is a text similarity between character strings of URLs (Uniform Resource Locators) of the documents and that is an appearance of written characters of URLs, the number of links to the document, and the number of links from the document, and

selecting the document based on the document type, and

wherein the creating creates the link list for the selected document based on the link importance and the particular characteristic of the character string of the URL, and

wherein said link importance being based on an inverse URL similarity and the link relation of the document, so that a link weight increases as URL similarity decreases.

30. (cancelled)

31. (currently amended) A computer readable record medium for recording a program that causes a computer to execute a process for creating a link list for a document group having a link relation, the program comprising:

collecting documents from a network;

assigning a link importance which indicates importance of the document to each document based on the link relation, including:

calculating a URL similarity that is a text similarity of character strings of URLs (Uniform Resource Locators) that represent the location of the documents and that is an appearance of written characters of URLs, and

calculating the link importance based on the URL similarity and the link relation of the document with said link importance being based on an inverse URL similarity and the link relation of the document, so that a link weight increases as URL similarity decreases;

determining a URL having a particular characteristic of a character string from the URLs of documents; and

creating a link list for listing less than a predetermined number of links to the documents based on the link importance and the particular characteristic of the character string of the URL.

32. (currently amended) A document searching method for searching documents linked by Uniform Resource Locators (URLs), comprising:

assigning a link importance to the document by determining a text similarity between URLs to and from a document with said link importance being based on an inverse URL

B1
similarity and the link relation of the document, so that a link weight increases as URL similarity decreases; and
accessing the document based on the link importance.

B2
33. (new) A document searching apparatus for searching a document group having a link relation for a document, comprising:
a link importance assigning unit weighting the link relation and assigning link importance which indicates importance of the document based on the weighted link relation to the document, said link importance assigning unit comprising a similarity calculating unit calculating a URL similarity that is a similarity URLs (Uniform Resource Locators) that represent the location of the documents and that is an appearance of written characters of URLs and said link importance assigning unit calculates the link importance based on the URL similarity and the link relation of the document; and
an accessing unit accessing the document based on the link importance.

34. (new) A document searching method for searching a document group having a link relation for a document, comprising:
assigning a link importance as an importance of the document calculated with weighting the link relation to the document, comprising:
calculating a similarity that is a similarity of URLs (Uniform Resource Locators) that represent the location of the documents; and
calculating the link importance based on the URL similarity and the link relation of the document; and
accessing the document based on the link importance.

35. (new) A document searching method for searching documents linked by Uniform Resource Locators (URLs), comprising:
assigning link importances to linked documents by determining a similarity between URLs to and from the documents; and
ranking and accessing the documents based on the link importances.
